Explosive Convexity of the Value Function in Learning Models

Axel Anderson† Lones Smith‡

Georgetown University University of Michigan
Economics Department Economics Department
Washington, DC 20057 Ann Arbor, MI 48109-1220

May 21, 2006

Abstract

We explore the behavior of value functions in binary state experimentation problems depending on the belief $x$ in the high state. We show that the nonlinear portion is proportional to $x^{\alpha_0}$ near $x = 0$ and $(1 - x)^{\alpha_1}$ near $x = 1$, once the discount factor $\delta$ passes a threshold. We provide explicit formulas for $\alpha_0$ and $\alpha_1 \in (1, 2)$. So if $v''$ exists near the extremes 0 and 1, it is unbounded. Each $\alpha_i$ is falling in the Kullback-Leibler distance between the two signal densities, given the statically optimal action. Implications of this finding include: (a) complete learning obtains above a known lower bound for $\delta$; (b) the value of additional small signals is of the order $x^{\alpha_0}$ and $(1 - x)^{\alpha_1}$ near $x = 0$ and $x = 1$. 

*Lones wishes to thank the NSF for continual funding.
†Email: aza@georgetown.edu
‡Email: lones@umich.edu; web: www.umich.edu/~lones
1 Introduction

This paper studies infinite horizon discounted dynamic Bayesian learning models. Value functions used in macroeconomics are generally concave functions of the state variable capital stock. But in a Bayesian learning context, they are convex in the state variable — the probability measure over a state of the world. This may well be our only stylized insight about Bayesian value functions. This paper sharpens our analytic understanding of the convex character of the value function. Our discovery, which struck us as surprising, concerns a special explosive character of the convexity.

We focus on the simple two state world so common in applications, where the state variable is the scalar belief $x$ in the high state. Here, it is well known that the value function flattens out as the decision maker grows increasingly patient: The marginal value of additional signals vanishes, and, intuitively, so does the convexity. Or does it? When the action space is the continuum, this paper finds otherwise. Once the discount factor passes a threshold below one — and possibly near zero — the second derivative of the value function, when it exists, explodes near any fixed point $x^*$ of the dynamical system. Differentiability aside, the nonlinear portion of the value is proportional to $(x - x^*)^\alpha$, for some $\alpha$ in $(1, 2)$. The second derivative quite clearly explodes near $x^*$.

This paper develops this result using a nonstandard functional equations analysis of (a) the policy equation for twice smooth policy functions in Proposition 1, and then (b) the Bellman equation in Proposition 2. We provide a relatively simple implicit formula for $\alpha$ that depends only on the signal structure and the expected stage payoffs. While our result is derived using functional equations, we do relate it to the informational foundations. Precisely, for sufficiently patient decision makers, we show in Proposition 3 that $\alpha$ is increasing in the (asymmetric) classical Kullback-Leibler “distance” between the state probability densities and the optimal action chosen for that state.

Proposition 1 is useful as part of a guess-and-verify methodology: When the map from states to actions is known, compute the implied asymptotic expression for the value function, and then verify whether the assumed map is optimal for this functional form. This was the counterfactual approach we adopted in Anderson and Smith (2004) to establish the impossibility of assortative matching in a model with learning.

We also show that it affords useful insights into complete learning and the value of informative signals. It turns out that this asymptotic form is inconsistent with

1
an interior fixed point (Proposition 4), where learning grinds to a halt. The reason? Given explosive convexity, the decision-maker become risk-loving near the extremes. The marginal informational gain to a small change in the statically optimal policy must swamp the static losses from such a change. Thus, complete learning obtains above this threshold. This lower bound is expressed in terms of the information process alone.

We next ask what is the value of an additional informative signal. It is obvious that this vanishes as we approach the extremes 0 and 1, since these are fixed points of Bayes rule. One might think them of order $x$, but in fact, we show in Proposition 5 that their order is larger — for instance, proportional to $x^\alpha$ near the belief $x = 0$.

Our proof is also instructive. Standard proofs that dynamic programming preserves a property appeal to the contraction property of the Bellman operator. The key to our characterization in Proposition 1 is its monotonicity. We construct two functions, both with the same asymptotic form. The Bellman operator pushes one of these functions down, and the other up. Ipso facto, the unique fixed point — namely, our value function — lies between these two functions, and so shares the same asymptotic form.

Our model is a special case of Easley and Kiefer (1988), restricted to two states. This is a standard special case in many experimentation models, and has been applied in many papers, such as McLennan (1984), Mirman, Samuelson, and Urbano (1993), Rustichini and Wolinsky (1995), Bergemann and Valimaki (2000), and recently Smith and Sorensen (2005). While it has been known that learning may be complete for sufficiently patient decision makers, we are the first to characterize the critical discount factor lower bound in a general learning problem.

A motivational example that captures our basic idea is found in §2. The model is laid out in §3. In §4, we analyze the functional equation for a fixed mapping from states to actions. We characterize the asymptotic form of the value function near any fixed point in §5. In §6, we present some applications of our characterization result.

2 A Motivational Example

We will ultimately consider two state learning problems, which admit a scalar belief state variable. Once we fix an optimal action in every state, the Bellman equation reduces to a simpler functional equation solved by the value function. We now abstract
from our model and construct a simple toy example that nevertheless affords intuition for our main results. Assume that a value function obeys a functional equation of the form:

\[ v(x) = (1 - \delta)x^2 + \delta \left[ \frac{1}{2}v(3x/2) + \frac{1}{2}v(x/2) \right] \tag{1} \]

namely, the weighted sum of a stage reward \( x^2 \) and an expected continuation value for a state variable \( x \), equally likely to rise to \( 3x/2 \) and fall to \( x/2 \) (thus a martingale). To admit this simple structure, we assume \( x \) positive (and so not a probability). Let us focus on the behavior of the solution in a neighborhood of the fixed point \( x = 0 \).

As a functional equation, (1) admits a homogeneous and particular solution. The particular solution is clearly quadratic. As with difference equations, the homogeneous solution assumes the geometric form \( x^\alpha \). Thus, the general solution is

\[ v(x) = \text{particular solution} + \text{homogeneous solution} = \frac{1 - \delta}{1 - 5\delta/4}x^2 + cx^\alpha \tag{2} \]

where \( \alpha \) solves

\[ \delta \left[ \frac{1}{2}(3/2)^\alpha + \frac{1}{2}(1/2)^\alpha \right] = 1. \tag{3} \]

Observe that the second derivative \( v'' \) is unbounded near 0 precisely when \( \alpha < 2 \). Is this true? By the power mean inequality (a generalization of arithmetic-geometric mean inequality), the left side of (3) exceeds \( \delta \) exactly when \( \alpha > 1 \). It is also a continuously and unboundedly increasing function of \( \alpha \) — for \((3/2)^\alpha\) rises faster than \((1/2)^\alpha\) falls, as differentiation reveals. The solution \( \alpha_\delta \) of (3) therefore uniquely exists, and obeys \( \alpha_1 = 1 \) and \( \alpha_\delta > 1 \) for any \( \delta < 1 \). Moreover, \( \alpha_\delta \) continuously falls in \( \delta \), and so there exists a unique \( \delta^* < 1 \) with \( \alpha_{\delta^*} = 2 \). One can easily verify here that in fact \( \delta^* = 4/5 \).

### 3 The Model

Consider an infinite horizon, discrete time decision problem, with an unobserved state \( \theta = H \) or \( \theta = L \). Assume that the decision maker learns about this state solely from observing payoffs. We focus on a two state world, so that the state variable are the scalar belief \( x \) that \( \theta = H \).

In each period, the decision maker chooses an action \( a \in [0, 1] \). He then receives a payoff \( \sigma \) from a compact set \( \Sigma \), according to the state contingent density over payoffs,
We assume for simplicity that the density $f^\theta$ is continuous in $a$. Upon observing the payoff, the decision maker updates his belief according to Bayes’ Rule.\footnote{The simplifying assumption that payoffs are the only observed signal is not necessary. As written our assumptions on $f^\theta$ carry through to assumptions on $\pi(x, a)$. We could complicate the analysis instead separate signals and payoffs and make similar assumptions directly on state contingent expected payoffs, at the cost of extra notation.} Thus, the updated belief that $\theta = H$ given the current belief, signal, and the action taken is $z(\sigma, x, a) \equiv f^H(\sigma, a)x/f(\sigma, x, a)$, where $f(\sigma, x, a) \equiv f^H(\sigma, a)x + f^L(\sigma, a)(1-x)$.

The decision maker is risk neutral and discounts the future with discount factor $0 \leq \delta < 1$. Define the expected payoff conditional on the action chosen as:

$$\pi(x, a) \equiv x \int \sigma f^H(\sigma, a)d\sigma + (1-x) \int \sigma f^L(\sigma, a)d\sigma.$$

A policy function is a map from states to actions: $a : [0, 1] \mapsto [0, 1]$, and the decision maker chooses a sequence of such policy functions, $\{a_t(\cdot)\}_{t=0}^\infty$ to maximize the expected discounted value of payoffs, $E[(1-\delta) \sum_{t=0}^\infty \delta^t \pi(x_t, a_t(x_t))]|x_0]$.

Following Easley and Kiefer (1988) the decision maker’s value function, $v$, uniquely solves the following Bellman equation:\footnote{The max exists as $\pi, z$, and $f$ are continuous, and $a$ is being chosen on a compact set.}

$$v(x) = \max_a \{(1-\delta)\pi(x, a) + \delta \int v(z(\sigma, x, a))f(\sigma, x, a)d\sigma\} \quad (4)$$

Note that $v$ must be convex by Easley and Kiefer (1988).

**Some Examples.** Our model applies to a wide range of two state learning problems covered in the literature, including:

**McLennan’s Ignorant Monopolist.** A monopolist is unsure about the demand curve it faces. In each period the monopolist sets a price and stochastic demand is realized. The conditional distribution of demand depends on the state of the world, and thus observed demand is a signal of the state of the world.

**Herding.** Smith and Sorensen showed that multi agent herding models can be reinterpreted as single agent learning problems. Our model covers this reinterpreted herding model.

**Learning by Doing Production.** Reduced form learning by doing would not be covered, but we would cover many learning by doing models based on primitives. For
Figure 1: $v_\beta$, $v'_\beta$, and $v''_\beta$. This graph illustrates $\lim_{x \to 0} v''_\beta(x) = \infty$. (Note $\beta_H > \beta_L$.)

example, if we let stochastic output be conditional on the unknown state of the world and let the firm chooses inputs in every period.

**Matching with Learning.** Agent’s with unknown types match in pairs to produce stochastic output. Each agent’s value in the market is a function of the public belief about their type.

As stated so far, our model would also cover optimal stopping problems with learning (ex. Bandit problems). However, our later assumptions will necessarily rule out this class of examples. The reason? Our techniques require that the same asymptotic expression hold for all posteriors when beliefs are suitably close to a fixed point. Typically in bandit problems when beliefs cross a switching threshold there is a discontinuous change in the convexity of the value function.

## 4 The Analysis with a Fixed Policy Function

Perhaps the easiest way to build intuition for our result is by approaching our decision problem as a functional equation. Specifically, if we were to substitute an optimal policy into the Bellman equation (4), the value function would be a solution of the resulting functional equation. More generally, we can consider the functional equation that results when we substitute any policy function into the Bellman equation. In this section we do so, and show that our asymptotics near a fixed point are driven by the solution of the functional equation, following from the martingale property of beliefs. and Jensen’s Inequality

In addition to it’s role as a pedagogical tool, the results in this section are independently useful in application. For example, proving that a particular policy is not
optimal can be of independent interest as a stand alone result or prove useful as a chunk of a larger proof. To use our results in this way one would assume a particular policy function, derive the asymptotic approximation associated with that policy, and then show that given this form, the posited policy function cannot be optimal. We provide an application following our results in which this approach proved fruitful.

Let \( a(\cdot) \) be a policy function. Assuming we fix such a policy function, the expected discounted value function satisfies the functional equation:

\[
v(x) = (1 - \delta)\pi(x, a(x)) + \delta \int v(\sigma(x, a(x))) f(x, a(x)) d\sigma \quad (5)
\]

The \( v \) that solves the above will be the value function, if and only if \( a(\cdot) \) solves the Bellman equation, but we shall abuse language and refer to the solution of (5) as the value function in this section for any arbitrary \( a(\cdot) \). Fixing a policy function also determines the stochastic evolution of beliefs (the belief process): \( x \mapsto z(\sigma, x, a(x)) \) with state contingent chance \( f^{\theta}(\sigma, a(x)) \). Beliefs will be stationary iff \( z(\sigma, x, a(x)) = x \) almost surely (a.s.). When \( x \in (0, 1) \) this is true iff \( f^H(\sigma, a) = f^L(\sigma, a) \) a.s. So, given policy \( a(x) \), \( x^* \) is a fixed point of the belief process iff \( x^* \in \{0, 1\} \) or \( f^H(\sigma, a(x^*)) = f^L(\sigma, a(x^*)) \) almost surely.

Rather than move immediately to our formal proof, we seek to build some intuition for why this functional form works, what determines the exponent \( \alpha \), and why \( \alpha \in (1, 2) \) for high \( \delta \). Toward this end, suspend disbelief and assume \( v(x) \) is proportional to a linear function plus \( (x - x^*)^\alpha \) near any fixed point of the belief process \( x^* \). Loosely, for this to work we need the expectation operator in (5) to preserve this functional form near \( x = x^* \), i.e.

\[
(x - x^*)^\alpha = \delta \int (z(\sigma, x, a(x)) - x^*)^\alpha f(x, a(x)) d\sigma \quad \text{(near } x = x^*)
\]

That is we need the right hand side to be proportional to \( (x - x^*)^\alpha \) and \( x \to x^* \), or,

\[
1 = \delta \int \left( \frac{z(\sigma, x, a(x)) - x^*}{x - x^*} \right)^\alpha f(x, a(x)) d\sigma \equiv \delta \Phi(\alpha, x, x^*) \quad \text{(as } x \to x^*)
\]

This intuitive approach suggests immediately that our approximation will fail unless \( \lim_{x \to x^*} \Phi(\alpha, x, x^*) \) exists and is bounded. In order for this limit to exist we need \( a(x) \)
to converge.

**Assumption 1** The policy function converges: \( \lim_{x \to x^*} a(x) \) exists.

If \( x^* \in \{0, 1\} \) this is all that is required for the requisite limit to boundedly exist, otherwise we need additional assumptions. The reason is that we need to use L’Hopital’s Lemma and thus must have differentiability, plus convergence.

**Assumption 2 (Differentiability)** The signal densities are differentiable in \( a \) at \( a = a(x^*) \), the policy function \( a \) is differentiable at \( x = x^* \), and \( \lim_{x \to x^*} a'(x) \neq 0 \) exists, where \( x^* \) is a fixed point of the belief process induced by \( a(\cdot) \).

**Lemma 1** Let \( x^* \) be a fixed point of the belief process induced by \( a(\cdot) \). If Assumption 1 and

- \( x^* \in \{0, 1\} \); or
- Assumption 2 holds,

then \( \lim_{x \to x^*} \Phi(\alpha, x, x^*) \) exists and is bounded for all \( 0 < \alpha < \infty \).

**Proof:** If \( \lim_{x \to x^*} \left[ (z(\sigma, x, a(x)) - x^*)/(x - x^*) \right] \) exists and is bounded for all \( \sigma \), then we are done. We have:

\[
\frac{z(\sigma, x, a(x)) - x^*}{x - x^*} = \frac{f_H(\sigma, a(x))x - f(\sigma, x, a(x))x^*}{f(\sigma, x, a(x))(x - x^*)} = \frac{f(x, x, a(x))(x - x^*) + f_H(\sigma, a(x)) - f_L(\sigma, a(x))}{f(\sigma, x, a(x))(x - x^*)} = 1 + \frac{x(1 - x)(f_H(\sigma, a(x)) - f_L(\sigma, a(x)))}{f(\sigma, x, a(x))(x - x^*)}
\]

If \( x^* \in \{0, 1\} \), then:

\[
\lim_{x \to x^*} \frac{z(\sigma, x, a(x)) - x^*}{x - x^*} = 1 + \frac{f_H(\sigma, a(x^*)) - f_L(\sigma, a(x^*))}{f(\sigma, x^*, a(x^*))} \lim_{x \to x^*} \frac{x(1 - x)}{x - x^*} = 1 + \frac{f^H(\sigma, a(x^*)) - f^L(\sigma, a(x^*))}{f(\sigma, x^*, a(x^*))}(1 - 2x^*)
\]

If instead, \( x^* \in (0, 1) \), then

\[
\lim_{x \to x^*} \frac{z(\sigma, x, a(x)) - x^*}{x - x^*} = 1 + \frac{x^*(1 - x^*)}{f(\sigma, x^*, a(x^*))} \lim_{x \to x^*} \frac{f_H(\sigma, a(x)) - f_L(\sigma, a(x))}{x - x^*}
\]

7
Note that since $x^*$ is an interior fixed point of the belief process, $f^H(\sigma, a^*(x^*)) = f^L(\sigma, a^*(x^*))$ almost surely, thus we again need to use L’Hopital to determine this limit. However, to do so requires that we take a derivative of $f^\theta$ and $a$ at $x^*$, which is precisely why we require Assumption 2 for the interior fixed point analysis:

$$\lim_{x \to x^*} \frac{f^H(\sigma, a(x)) - f^L(\sigma, a(x))}{x - x^*} = (f_a^H(\sigma, a(x^*)) - f_a^L(\sigma, a(x^*)))a'(x^*)$$

□

We have established that $\Phi(\alpha, x^*) \equiv \lim_{x \to x^*} \Phi(\alpha, x, x^*)$ is well defined and bounded, but we need more for equation (6) to hold. We need given $\delta$ there exists an $\alpha_\delta$ that solves:

$$1 = \delta \Phi(\alpha_\delta, x^*) \quad (7)$$

Additionally, we would like to characterize the $\alpha_\delta$ that solves this equation. For interior fixed points ($x^* \in (0, 1)$) we need one more assumption to do so.

**Assumption 3 (Local Uniqueness)** Uninformative actions are locally unique. If $f^H(\sigma, a) = f^L(\sigma, a)$ a.s. then $\exists \varepsilon > 0$ such that $f^H(\sigma, \hat{a}) \neq f^L(\sigma, \hat{a})$ on a set of positive measure $\forall |a - \hat{a}| \leq \varepsilon$.

Now we are ready to characterize $\Phi$, the key to our approximation.

**Lemma 2** Let $x^*$ be a fixed point of the belief process induced by $a(\cdot)$ then if Assumption 1 holds:

1. $\Phi(\alpha, x^*)$ is continuous in $\alpha$;
2. $\Phi(1, x^*) = 1$ and $\Phi(2, x^*) > 1$;
3. $\Phi(\alpha, x^*)$ and $\Phi(\alpha, x, x^*)$ strictly increase in $\alpha$.

If either $x^* \in \{0, 1\}$ or Assumption 2 and Assumption 3 hold.

**Proof of Part 1:** Immediately follows from Lemma 1.

**Proof of Part 2:** The first claim in part 2 follows from the Martingale property of beliefs, i.e. that

$$\int (z(\sigma, x, a(x)) - x^*)f(\sigma, x, a(x))d\sigma = x - x^* \quad \forall \ x$$
For the second claim in part 2, first consider the case $x^* = 0$ ($x^* = 1$ is similar):

\[
\Phi(2, 0) = \lim_{x \to 0} \int \frac{\int H(\sigma, a(x))^2 x^2}{f(\sigma, x, a(x))} f(\sigma, x, a(x)) d\sigma = \int \frac{\int H(\sigma, a(0))^2}{f L(\sigma, a(0))} d\sigma > 1
\]

To see this last inequality, write

\[
\int \frac{\int H(\sigma, a(0))^2}{f L(\sigma, a(0))} d\sigma - 1 = \int \frac{\int H(\sigma, a(0))}{f L(\sigma, a(0))} \left[ H(\sigma, a(0)) - f L(\sigma, a(0)) \right] d\sigma > 0
\]

since $f H \neq f L$, $f H, f L \geq 0$ and $\int f H(\sigma) d\sigma = \int f L(\sigma) d\sigma = 1$.

For the $x^* \in (0, 1)$ case, from Lemma 1 we have:

\[
\Phi(2, x^*) = \int \left[ 1 + \frac{x^*(1 - x^*)}{\int f(\sigma, x^*, a(x^*))} \left( f H(\sigma, a(x^*)) - f L(\sigma, a(x^*)) a'(x^*) \right)^2 f(\sigma, x^*, a(x^*)) d\sigma = 1 + [x^*(1 - x^*)]^2 \int \frac{\left[ f H(\sigma, a(x^*)) - f L(\sigma, a(x^*)) a'(x^*) \right]^2}{f(\sigma, x^*, a(x^*))} d\sigma
\]

where $\int (f H(\sigma, a(x^*)) - f L(\sigma, a(x^*)) d\sigma = 0$ follows from $\int [f H(\sigma, a) - f L(\sigma, a)] d\sigma = 0$ for all $a$. Notice that all potentially non positive terms in the integrand are squared. Thus, our result only fails to hold if the integrand is zero almost everywhere. Since $x^* \in (0, 1)$ we have $f H(\sigma, a(x^*)) - f L(\sigma, a(x^*)) = 0$ for all $\sigma$. Thus, the derivative of this difference cannot be zero as well or we would violate Differentiability and Local Uniqueness.

**Proof of Part 3:** By the Power Mean Inequality $M_\theta(x) = (\int \omega(\sigma)y(\sigma)^\theta d\sigma)^{1/\theta}$ is rising in $\theta$ if $\int \omega(\sigma)d\sigma = 1$. Setting $\theta = 1/\alpha$, $\omega(\sigma) = f(\sigma, x^*, a(x^*))$ and $y(\sigma) = [z(\sigma, x^*, a(x^*)) - x^*]/[x - x^*]$ yields $\Phi(x^*)^{1/\alpha}$ increasing in $\alpha$. We claim $\Phi(x^*) \geq 1$ for all $\alpha > 1$. To see this, assume instead that $\Phi(x^*) < 1$ for some $\alpha > 1$. This implies $\Phi(x^*)^{1/\alpha} < 1$. Combining this with the Power Mean Inequality and that $1 = \Phi(1, x^*)$ yields:

\[
1 = \Phi(1, x^*) < \Phi(\alpha, x^*)^{\frac{\alpha}{\alpha}} < 1
\]

a contradiction. Now to finish the proof, assume that there exists $1 < \alpha < \bar{\alpha}$ such that:

$\Phi(\bar{\alpha}, x^*) < \Phi(\alpha, x^*)$, the Power Mean Inequality then yields:

\[
\Phi(\bar{\alpha}, x^*)^{\frac{\alpha}{\alpha}} \geq \Phi(\alpha, x^*)^{\frac{1}{\alpha}} \Rightarrow \Phi(\bar{\alpha}, x^*) \geq \Phi(\alpha, x^*)^{\frac{\alpha}{\alpha}} \geq \Phi(\alpha, x^*)
\]
where the last inequality follows from \( \hat{\alpha}/\alpha > 1 \) and \( \Phi(\alpha, x^*) \geq 1 \). The proof that \( \Phi(\alpha, x, x^*) \) is increasing in \( \alpha \) is analogous.

\[ \square \]

**Lemma 3** Let \( x^* \) be a fixed point of the belief process induced by \( a(x) \) and \( \alpha_\delta \) be defined by equation 7, then \( \alpha_\delta \in (1, 2) \) is decreasing in \( \delta \) for all \( \delta > \delta^* \) if the requirements of Lemma 2 are satisfied, where \( \delta^* \) solves \( 1 \equiv \delta^* \Phi(2, x^*) \).

**Proof:** Existence of an \( \alpha_\delta \) that solves (7) follows from Lemma 2. Uniqueness and the characterization of the threshold \( \delta \) follows from part 3.

Now we are ready to establish our main proposition for fixed policy functions. For this proposition we need a rather strong assumption about the expected flow payoff \( \pi(x, a(x)) \).

**Assumption 4** The expected flow payoff function, \( \pi(x, a(x)) \) is strictly convex in \( x \).

If \( a \) is constant \( \pi(x, a) \) is linear. If instead \( a \) is the statically optimal policy, \( \pi(x, a(x)) \) must be weakly convex. In addition, for such statically optimal policies, strict convexity obtains at any \( x \) iff we cannot choose a statically optimal policy that is constant at that \( x \). Note that the utility of our fixed policy results comes from guess and verify techniques: choose a candidate \( a(x) \) and then verify that the asymptotic form this implies either satisfies or violates the Bellman Equation. One likely candidate would be the statically optimal policy. Note that we substantially weaken this assumption when \( a(x) \) is an optimal policy.

**Assumption 5** There exists a linear function, \( L(x) \), such that \( \pi(x, a(x)) - L(x) = O((x - x^*)^2) \) as \( (x \rightarrow x^*) \).\(^3\)

Note that a second order Taylor Series approximation to \( \pi(x, a(x)) \) would satisfy this assumption.

**Proposition 1** Fix a policy function \( a(\cdot) \) and let Assumptions 1, 4, and 5 hold and define \( 1 \equiv \delta^* \Phi(2, x^*) \). Then \( \forall \delta > \delta^* \) there exists \( c > 0 \), such that the value function is proportional to a linear function plus \( c(x - x^*)^{\alpha_\delta} \), where \( \alpha_\delta \in (1, 2) \) solves (7) as long as \( x^* \in \{0, 1\} \) or Assumptions 2 and 3 hold. Further if \( f^0 \) and \( a(\cdot) \) are \( C^2 \) then the second derivative exists, is continuous, and is proportional to \( c(\alpha_\delta - 1)\alpha_\delta(x - x^*)^{\alpha_\delta - 2} \).

\(^3\)Notational reminder: \( g(x) = O(h(x)) \) iff \( g(x)/h(x) \) converges to a constant, while \( g(x) = h(x)(1 + o(1)) \) \( (x \rightarrow x^*) \) iff \( g(x) \sim h(x) \) near \( x^* \) iff \( \lim_{x \rightarrow x^*} g(x)/h(x) = 1 \).
Proof: Define the operator: \( T : \mathcal{L}_\infty \to \mathcal{L}_\infty \) as:

\[
Tv(x) = (1 - \delta)\pi(x, a(x)) + \delta \int v(z(x, a(x)))f(\sigma, x, a(x))d\sigma
\]

Notice that this is a monotonic contraction in the sup norm, so that the solution is unique and if we show \( Th \geq h \) for some bounded function \( h \), then we know our unique solution \( v \) must satisfy \( v \geq h \). We shall construct two bounded functions: \( \underline{v} \) and \( \bar{v} \), both of which have the correct asymptotic form and then prove that \( T\underline{v} \geq \underline{v} \) and \( T\bar{v} \leq \bar{v} \). This then implies that \( \underline{v} \leq v \leq \bar{v} \), which in turn implies that \( v \) has the correct asymptotic form. In particular, we show there exists a linear function \( L(x) \) and constants \( A, B, C > 0 \) such that:

\[
\begin{align*}
\text{Claim 1: } \underline{v}(x) &= L(x) + A \left[ (x - x^*)^{\alpha\delta} + (x - x^*)^2 \right] \Rightarrow T\underline{v} \geq \underline{v} \\
\text{Claim 2: } \bar{v}(x) &= L(x) + B \left[ (x - x^*)^{\alpha\delta} - (x - x^*)^2 \right] - C(x - x^*)^{\beta + 1} \Rightarrow T\bar{v} \leq \bar{v}
\end{align*}
\]

where \( \beta \in \arg\max_x \beta(x) \) s.t. \( \delta \Phi(\beta(x), x, x^*) = 1 \).

Since \( \pi(x, a(x)) \) is strictly convex, we can construct a linear function \( L(x) \) with the following properties:

- \( \pi(x^*, a(x^*)) - L(x^*) = 0 \)
- \( \pi(x, a(x)) - L(x) \geq 0 \) \( \forall x \)
- \( \forall \varepsilon > 0 \exists M_\varepsilon \) s.t. \( \pi(x, a(x)) - L(x) \geq M_\varepsilon \) \( \forall |x - x^*| \geq \varepsilon \)

**Proof of Claim 1:** Substituting the given \( \underline{v} \) into our operator and manipulating we have: \( T\underline{v} \geq \underline{v} \iff \)

\[
(1 - \delta)[\pi(x, a(x)) - L(x)] \geq A(x - x^*)^{\alpha\delta}[1 - \delta \Phi(\alpha, x, x^*)] + A(x - x^*)^2[1 - \delta \Phi(2, x, x^*)]
\]

By construction both sides of this inequality are zero at \( x = x^* \).

Since \( \Phi(\alpha, x, x^*) \) is continuous in \( x \) near \( x = x^* \) and \( 1 = \delta \tilde{\Phi}(\alpha, x^*) \leq \Phi(2, x^*) \) by Lemma 2, we have that:

\[
\begin{align*}
(x - x^*)^{\alpha\delta}(1 - \Phi(\alpha, x^*)) &= O((x - x^*)^{\alpha\delta + 1}) \\
(x - x^*)^2(1 - \Phi(2, x^*)) &= O((x - x^*)^2)
\end{align*}
\]
So the second term on the right hand side is dominant, further since \( 1 - \delta \Phi(2, x^*) < 0 \), there exists an \( \varepsilon > 0 \) such that the right hand side is negative for all \( |x - x^*| \leq 0 \). Note that this \( \varepsilon \) is critically not a function of \( A \). Thus, the inequality holds in a neighborhood of \( x^* \) \( \forall A > 0 \). We have shown that \( \Phi(\alpha, x, x^*) \geq 1 \) for all \( \alpha > 1 \) (Jensen’s Inequality plus the Martingale property of beliefs). This implies that the right hand side of (8) does not exceed \( A(1 - \delta) \), thus (8) is satisfied whenever:

\[
\pi(x, a(x)) - L(x) \geq A
\]

Finally since \( \exists a \) \( M_{\varepsilon} > 0 \) such that \( \pi(x, a(x)) - L(x) \geq M_{\varepsilon} \) \( \forall |x - x^*| \geq \varepsilon \), by setting \( A = M_{\varepsilon} \), Claim 1 is established.

**Proof of Claim 2:** Substituting the given \( \bar{v} \) into our operator and manipulating we have:

\[
(1 - \delta)[\pi(x, a(x)) - L(x)] \leq B(x - x^*)^{\alpha_2}[1 - \delta \Phi(\alpha_3, x, x^*)] - B(x - x^*)^2[1 - \delta \Phi(2, x, x^*)]
\]

\[
-C(x - x^*)^{\beta+1}[1 - \delta \Phi(\beta+1, x, x^*)]
\]

By construction both sides of this inequality are zero at \( x = x^* \).

Also, by construction \(-[1 - \delta \Phi(\beta+1, x, x^*)] > 0 \). Further \( 2 < \beta+1 \), thus near \( x = x^* \) the behavior of the RHS of this inequality is determined by \(-B(x - x^*)^2[1 - \delta \Phi(2, x, x^*)]\), which converges to a positive constant. By assumption, \( \pi(x, a(x)) - L(x) \) is \( O((x - x^*)^2) \), thus for \( B \) large enough the inequality will hold near \( x = x^* \) for any \( C \geq 0 \). Let \( \varepsilon > 0 \) be such that the inequality holds \( \forall |x - x^*| \leq \varepsilon \) when \( C = 0 \). (since \(-[1 - \delta \Phi(\beta+1, x, x^*)] \geq 0 \), the inequality will hold on \([0, \varepsilon] \) for all \( C \geq 0 \).)

We have \(-[1 - \delta \Phi(\beta, x, x^*)] \geq 0 \). Thus, since \( \Phi \) is strictly increasing in it’s first argument, there exists \( M > 0 \) such that \(-[1 - \delta \Phi(\beta+1, x, x^*)] \geq M \). Thus,

\[
-C(x - x^*)^{\beta+1}[1 - \delta \Phi(\beta+1, x, x^*)] \geq C\varepsilon^{\beta+1}M \ \forall |x - x^*| \geq \varepsilon
\]

Since, \( \pi(x, a(x)) - L(x) \) is bounded we can choose \( C \) large enough such that inequality (9) hold for all \( |x - x^*| \geq \varepsilon \). Thus, \( v \) has the form given.

We have found the tail asymptotic order on the value function. This seems tantalizingly close to our goal of \( v''(x) \propto (x - x^*)^{\alpha-2} \). But whereas integrating asymptotic
relations is a fully valid exercise, differentiation requires regularity conditions on the derivative (results known as ‘Tauberian Theorems’). Fortunately, by §7.3 in De Bruijn (1958), monotonicity is one such condition for our context, with an implicitly defined function.

**Step 1** Near \( x = 0 \), \( v'(x) \) exists, is continuous, and \( v'(x) \sim v'(x^*) + c \alpha \delta(x-x^*)^{\alpha-1} \).

*Proof:* Our operator, \( T \), preserves convexity and \( \pi(x, a(x)) \) is convex, thus \( \pi \) is convex. This implies that where it exists, the first derivative, \( v' \) is monotonic. Hence De Bruijn’s condition is met, provided \( v' \) exists near \( x = x^* \), as next established.

In fact, we claim, there exists \( \varepsilon > 0 \) such that \( v' \) exists and is continuous on \([0, \varepsilon]\). To simplify the exposition, abuse notation and define \( z(\sigma, x) \equiv z(\sigma, x, a(x)) \), \( f(\sigma, x) \equiv f(\sigma, x, a(x)) \), and \( \pi(x) \equiv \pi(x, a(x)) \). Thus, functions such as \( z_x(\sigma, x) \) are understood to mean \( z_2(\sigma, a(x)) + z_3(\sigma, a(x))a'(x) \). Differentiating equation (5) yields:

\[
v'(x) = \eta(v, x) + \delta \int v'(z(\sigma, x)) z_x(\sigma, x) f(\sigma, x) d\sigma
\]

where \( \eta(v, x) \equiv (1 - \delta)\pi'(x) + \delta \int v(z(\sigma, x)) f_x(\sigma, x) d\sigma \). From this we form the functional equation:

\[
\varphi'(x) = \eta(v, x) + \delta \int \varphi'(z(\sigma, x)) z_x(\sigma, x) f(\sigma, x) d\sigma.
\] (10)

Notice that we have used \( v \) as an argument in \( \eta \), while allowing \( \varphi' \) to vary. Clearly, \( \varphi' = v' \) solves this functional equation. Since \( \pi \) is \( C^1 \) and \( v \) is continuous, \( \eta \) is continuous. In addition, we have that \( \delta \int z_x(\sigma, x) f(\sigma, x) d\sigma < 1 \) near \( x = x^* \). To see this, recall \( \int z(\sigma, x) f(\sigma, x) = x \) for all \( x \), thus

\[
\int [z_x(\sigma, x) f(\sigma, x) + z(\sigma, x) f_x(\sigma, x)] d\sigma = 1
\]

and so,

\[
\lim_{x \to x^*} \int z_x(\sigma, x) f(\sigma, x) d\sigma = 1 - \lim_{x \to x^*} \int z(\sigma, x) f_x(\sigma, x) d\sigma = 1
\]

were the final equality follows from \( z(\sigma, x^*) = x^* \) and \( \int f_x(\sigma, x) d\sigma = 0 \). Finally, \( \delta \int z_x(\sigma, x) f(\sigma, x) d\sigma \) is continuous in \( x \). Altogether, Lemma 4 in Choczewski (1961) implies that the solution to (10) uniquely exists near \( x = x^* \) and is continuous. \( \Box \)
Now, being convex on \([0, 1]\), the policy value function is a.e. twice differentiable; however, the easy deduction of Step 1 is simply not an option for the second derivative, since \(v''\) is not monotonic. We proceed down a different route. Step 1 rules out any ‘kinks’ near 0, where the first derivative jumps up. The next result asserts that near 0, the second derivative is not merely nonnegative (when it exists), but is in fact locally continuous. Trivially, this implies that it everywhere exists near 0.

**Step 2** There exists \(\varepsilon > 0\) such that \(v''\) exists and is continuous on \((0, \varepsilon]\).

**Proof:** Twice differentiate the functional equation (5) and fix \(v, v'\) to form:

\[
\varphi''(x) = \gamma(v, v', x) + \delta \int \varphi''(z(\sigma, x))z_x(\sigma, x)^2f(\sigma, x)d\sigma
\]

where \(\gamma(v, v', x) \equiv (1-\delta)\pi''(x) + \delta \int [v(z(\sigma, x))f_{xx}(\sigma, x) + v'(z(\sigma, x))(z_{xx}(\sigma, x)f(\sigma, x) + 2z_x(\sigma, x)f_x(\sigma, x))]d\sigma\)

The continuity of \(\gamma\) follows from \(\pi\) and \(f\ C^2\) and the continuity of \(v\) and \(v'\). If we could show that \(\delta \int z_x(\sigma, x)^2f(\sigma, x)d\sigma < 1\) near \(x = x^*\) we could immediately again apply Lemma 4 in Choczewski (1961) and complete this step. However, this integral is actually greater than one near \(x = x^*\).

In order to complete our step we must normalize equation (11) in such a way that the resulting sum of coefficient is less than one. To do so, multiple both sides of (11) by \(x - x^*\) and define \(\phi''(x) = \varphi''(x)(x - x^*)\) to get:

\[
\phi''(x) = \gamma(v, v', x)(x - x^*) + \delta \int \phi''(z(\sigma, x))\left(\frac{x - x^*}{z(\sigma, x) - x^*}\right)z_x(\sigma, x)^2f(\sigma, x)d\sigma
\]

Notice that if we can show \(\phi''\) exists and is continuous on some interval \([0, \varepsilon]\) then \(\varphi''\) exists and is continuous on \((0, \varepsilon]\) as required. We showed in Lemma 1 that \(\lim_{x \to x^*} \frac{x - x^*}{z(\sigma, x) - x^*}\) boundedly exists and by L’Hopital equals \(1/z_x(\sigma, x)\). Thus,

\[
\lim_{x \to x^*} \int \left(\frac{x - x^*}{z(\sigma, x) - x^*}\right)z_x(\sigma, x)^2f(\sigma, x)d\sigma = \lim_{x \to x^*} \int z_x(\sigma, x)f(\sigma, x)d\sigma = 1
\]

where the last equality was proven in Step 1. \(\square\)
We know that $v''$ exists and is continuous and can now deduce our intuited asymptotic expression.

**Step 3** We have $v''(x) \sim c \alpha \delta (\alpha \delta - 1)(x - x^*)^{\alpha \delta - 2}$.

**Proof:** Define $\phi''_{\alpha \delta}(x) = v''(x)(x - x^*)^{2 - \alpha \delta}$. By Step 2, $\phi''_{\alpha \delta}$ is continuous near $x^*$. If $\phi''_{\alpha \delta}(x)$ converges to a finite constant $c$ as $x \to x^*$, we are done — for then $\phi''_{\alpha \delta}(x) \sim c$; this means that $v''(x) \sim c(x - x^*)^{\alpha \delta - 2}$. Otherwise, $\phi''_{\alpha \delta}(x)$ explodes near $x^*$, so that for all $M > 0$, there exists $\eta_M > 0$ with $(x - x^*)^{2 - \alpha \delta}v''(x) > M$ on $[0, \eta_M]$. So,

$$v'(\eta_M) - v'(x^*) = \int_{x^*}^{\eta_M} v''(t)dt \geq \int_{x^*}^{\eta_M} M(t - x^*)^{\alpha \delta - 2}dt = M(\eta_M - x^*)^{\alpha \delta - 1}/(\alpha \delta - 1).$$

Since $M$ is arbitrarily large for small $\eta_M$, this violates Step 1. Thus, $c < \infty$. □

Since $\alpha \in (1, 2)$, this means convexity is unbounded for low $x$. Why is convexity unbounded as $x \to 0$? One can show that $v'(1) - v'(0) = \int v''(x)dx$ is a constant in the discount factor. But, as is well known, the value function flattens out (becomes linear) as the discount factor converges to 1. To resolve this apparent conflict, convexity must accumulate near $x = 0$ and $x = 1$.

**Example: Matching with Learning.** Anderson and Smith (2004) consider a dynamic model of matching with learning. Specifically, agents match in pairs to produce stochastic output in each of an infinite number of periods. Each agent can be one of two unobservable types: $G$ or $B$. Let an agent’s reputation, $x$, be the publicly observable probability that they are type $G$.

Let $g^{\vartheta \vartheta'}(\sigma)$ be the chance that publicly observable output $\sigma$ is produced when types $\vartheta, \vartheta' \in \{G, B\}$ match. For simplicity assume matching is symmetric so that $g^{GB} = g^{BG}$. Assume no search frictions so that the decision problem facing each individual agent of reputation $x$ is to choose whom to match with to maximize his discounted stream of output shares given the market determined shares of output. Let $v(x)$ be an equilibrium value of a reputation $x$ in this market, which must solve:

$$v(x) = \max_a \{ (1 - \delta) \lambda(x, a)\pi(x, a) + \delta \int v(z(\sigma, x, a))f(\sigma, x, a)d\sigma \}$$
where \( \lambda(x,a) \) is the share of output \( x \) gets when matched with \( a \), and:

\[
 f^H(\sigma,a) \equiv ag^{GG}(\sigma) + (1-a)g^{GB}(\sigma) \quad f^L(\sigma,a) \equiv ag^{GB}(\sigma) + (1-a)g^{BB}(\sigma)
\]

In a static matching model without search frictions, like reputations would match in equilibrium if \( \pi_{xa} \geq 0 \). So now ask, when do like types match in this dynamic model? Assuming like types do match, that is \( a(x) = x \), the Bellman equation becomes:

\[
 v(x) = (1-\delta)\pi(x,x)/2 + \delta \int v(z(\sigma,x,x))f(\sigma,x,x)d\sigma
\]

Notice that \( \pi, z, \) and \( f^\theta \) are all \( C^\infty \). Assuming that \( G \) types have higher expected production than type Bs (i.e. \( \int \sigma g^{GG} > \int \sigma g^{GB} > \int \sigma g^{BB} \)), then \( \pi(x,x) \) will be strictly convex if \( \pi_{xa} \geq 0 \). Thus, we can apply Proposition 1 to approximate \( v \) near \( x = 0 \) and \( x = 1 \). For example, for some \( c > 0 \), near \( x = 0 \), \( v(x) = L(x) + cx^{\alpha_\delta}(1+o(1)) \), for all \( \delta > \delta^* \), where:

\[
 2L(x) \equiv \int \sigma g^{BB}(\sigma)d\sigma + x \int \sigma [g^{GB}(\sigma) - g^{BB}(\sigma)] d\sigma
\]

and \( \alpha_\delta \) and \( \delta^* \) solve:

\[
 1 \equiv \delta^* \int \frac{g^{BB}(\sigma)^2}{g^{GB}(\sigma)} d\sigma \quad 1 \equiv \int \left( \frac{g^{BB}(\sigma)}{g^{GB}(\sigma)} \right)^{\alpha_\delta} g^{GB}(\sigma)d\sigma
\]

Given these approximations one can then calculate whether agents with reputations near \( x = 0 \) and \( x = 1 \) would like to deviate from \( a(x) = x \). It turns out that it is almost surely the case that either (or both) very high or very low reputation agents will deviate from \( a(x) = x \). Thus, \( a(x) = x \) can never be an equilibrium matching pattern, despite types being complements in production.

## 5 The Value Function Near a Fixed Point

Since the optimal actions in states \( x = 0,1 \) is \( a^*(0), a^*(1) \), and the state is learned in the long-run as \( \delta \to 1 \), the limit value function is

\[
 \bar{v}(x) = (1-x)\pi(0,a^*(0)) + x\pi(1,a^*(1)) = A' + B'x
\]
In this section we present our main result, that for any fixed point $x^*$:

$$v(x) = A + Bx + C(x - x^*)^2 - \alpha (1 + o(1)) \quad (x \to x^*),$$

for some $C > 0$, where $\alpha \in (1, 2)$ for $\delta$ high enough. Indeed, in the $\delta = 1$ limit, we have

$$v(x) = A(x^*) + B(x^*)x + C(x^*)(x - x^*) = A' + B'x$$

and thus, equating coefficients, we have near $A(0) = A' + C(0)$ and $B(0) = B' - C(0)$, and $A(1) = A' + C(1)$ and $B(1) = B' - C(1)$.

### 5.1 Analysis

Let $a^* : [0, 1] \mapsto [0, 1]$ be a dynamically optimal policy, that is, solve the Bellman Equation. Let $x^*$ be a fixed point of the belief process induced by $a^*$. Define $a_0(x) \in \arg\max_a \pi(x, a)$. We replace some of the earlier assumptions as follows.

**Assumption 6** The statically optimal policy function converges: $\lim_{x \to x^*} a_0(x)$ exists.

Easley and Kiefer (1988) showed that the dynamically optimal policy, $a^*$, converges to the statically optimal policy, $a_0$, at any fixed point of the belief process induced by $a^*$. Thus, an immediate corollary of Assumption 6 is that $\lim_{x \to x^*} a^*(x)$ exists.

As in the fixed policy function case, we make strong assumptions when $x^* \in (0, 1)$.

**Assumption 7** (Differentiability) The signal densities are differentiable in $a$ at $a = a_0(x^*)$, the policy function $a^*$ is differentiable at $x = x^*$, and $\lim_{x \to x^*} (a^*)(x)$ exists.

Define $\Phi(\alpha, x, x^*)$ with respect to the dynamically optimal policy $a^*$. Note that $\alpha_\delta$ will be the same regardless of which policy is used since $a^*(x) \to a_0(x)$.

**Lemma 4** Let Assumption 6 hold and either: $x^* \in \{0, 1\}$ or Assumption 7 hold, then:

1. $\Phi(\alpha, x, x^*)$ increases in $\alpha$
2. $\Phi(\alpha, x^*)$ boundedly exists
3. There exists $\delta^* < 1$ such that $\alpha_\delta \in (1, 2)$ decreases in $\delta$ for all $\delta > \delta^*$. 

17
Proof: In the $x^* \in \{0, 1\}$ the fact that $a^*(x)$ converges to $a^0(x)$ plus Assumption 6 and Lemmas 2 and 3 establish Parts 1-3. For the $x^* \in (0, 1)$ case the Assumption 7 along with the same earlier Lemmas is sufficient. □

**Assumption 8 (Strict Convexity)** The statically optimal flow profit function, $\max_a \pi(x, a)$, is strictly convex

As mentioned previously, this function is always weakly convex. As long as the optimal policy changes in $x$, this function will be strictly convex. Also, once we assume strict convexity, we can define $L(x)$ such that $\max_a \pi(x, a) - L(x)$ is zero at $x^*$, nonnegative, and uniformly bounded below by some $M_\varepsilon$ for any $|x - x^*| \geq \varepsilon$. As before we need to make an assumption about the rate at which $\max_a \pi(x, a) - L(x)$ converges to zero.

**Assumption 9** Given $L(x)$ as defined above, $\max_a \pi(x, a) - L(x) = O((x - x^*)^2)$.

Note that if we can take a second order Taylor Series approximation to $\max_a \pi(x, a)$ then Assumption 9 is immediately satisfied. Since, $\pi(x, a)$ is continuous in $x$, $\max_a \pi(x, a)$ is continuous in $x$, but this is not necessarily sufficient for a second order Taylor series.

**Proposition 2** Let Assumptions 6, 8, and 9 hold and either $x^* \in \{0, 1\}$ or Assumption 7 hold, then there exists $c > 0$, such that for all $\delta \geq \delta^*$, the value function is proportional to a linear function plus $c(x - x^*)^{\alpha_\delta}$, with $\alpha_\delta \in (1, 2)$, where $\delta^*$ solves $1 \equiv \delta^* \Phi(2, x^*)$.

Proof: Clearly the optimal value function must weakly exceed the value obtained by fixing a statically optimal policy. We can then apply Proposition 1 to conclude that there exists $A > 0$ such that

$$v(x) \geq L(x) + A \left[ (x - x^*)^{\alpha_\delta} + (x - x^*)^2 \right]$$

so that our lower bound is established.

To establish an upper bound note that our value function $v$ must satisfy:

$$v(x) \leq (1 - \delta)\pi(x, a^0(x)) + \delta \int v(z(\sigma, x, a^*(x)))f(\sigma, x, a^*(x)))$$
Form the operator:

\[ \tilde{T}v(x) = (1 - \delta)\pi(x, a^0(x)) + \delta \int v(z(x, a^0(x))) f(\sigma(x, a^0(x))) \]

which is a monotonic contraction in the sup-norm, the fixed point of which must lie above our value function. Thus, if we can show that:

\[ \tilde{v}(x) = L(x) + B \left[ (x - x^*)^\alpha - (x - x^*)^2 \right] - C(x - x^*)^{\beta+1} \Rightarrow \tilde{T}\tilde{v} \leq \tilde{v} \]

where again the \( \Phi \) functions are defined with respect to \( a^* \), then we are done. Substitution yields the following equivalent inequality:

\[
(1 - \delta)[\pi(x, a^0(x)) - L(x)] \leq B(x - x^*)^{\alpha\delta}[1 - \delta \Phi(\alpha, x, x^*)] - B(x - x^*)^2[1 - \delta \Phi(2, x, x^*)] \]

\[
- C(x - x^*)^{\beta+1}[1 - \delta \Phi(\beta + 1, x, x^*)] \tag{12}
\]

which we have shown to be satisfied in the proof of Proposition 1, as long as \( \Phi(2, x, x^*) \) is continuous in \( x \) near \( x = x^* \). \( \square \)

**Example: The Ignorant Monopolist**  For simplicity assume that a monopolist’s demand can either be 0 or 1 in each period, i.e. \( \sigma \in \Sigma = \{0, 1\} \), and let \( f^\theta(a) \) the probability that demand is 1 given price \( a \). To avoid perfectly revealing prices assume that \( 0 < f^\theta(a) < 1 \) for all \( a \in [0, 1] \), and 0 otherwise. Thus, \( \pi(x, a) = a[x f^H(a) + (1 - x) f^L(a)] \).

Here there are fixed points at \( \{0, 1\} \) in addition to possible interior fixed points whenever \( f^H(a^0(x)) = f^L(a^0(x)) \) (essentially whenever expected demand curves cross). Thus, a sufficient condition for Local Uniqueness to obtain is that expected demand curves have different slopes at any crossing point, or weaker: that crossing points are locally unique. To understand the other assumptions define expected demand \( D(x, a) \equiv x f^H(a) + (1 - x) f^L(a) \), so that the static maximization problem is \( \max_a aD(x, a) \). If we assume \( f^\theta \) differentiable (necessary for our interior fixed point analysis), we have first order condition:

\[ 1 + \kappa(x, a^0(x)) = 0 \]

where \( \kappa \) is the elasticity of expected demand. Thus, we need \( \kappa \) well behaved to satisfy
our assumptions (convergent, and $C^1$ at $x^*$). For strict convexity of $\pi(x,a^0(x))$, it is sufficient that the elasticity of expected demand is not constant in $x$.

5.2 Kullback-Leibler Distance

We have established that the local curvature of the value function is determined by $\alpha_\delta$ near extremal fixed points. As we show below, for $\delta$ above our threshold, interior fixed points cannot exist, so we focus now on $x^* \in \{0,1\}$. In fact, in this section we assume $x^* = 0$ to conserve space. We have shown that $\alpha_\delta$ solves:

$$\delta \int \left( \frac{f^H(\sigma, a^*(0))}{f^L(\sigma, a^*(0))} \right)^{\alpha_\delta} f^L(\sigma, a^*(0)) \, d\sigma \equiv 1,$$

but have offered no interpretation for this formula. We can relate this formula to the information literature when $\delta$ is near 1.

The Kullback-Leibler distance between probability densities $a$ and $b$ is defined as:

$$KL(a,b) = \int a(\sigma) \log(a(\sigma)/b(\sigma)).$$

This is a commonly used statistic of two conditional signal distributions in the information literature outside of economics. It obeys $KL(a,b) \geq 0$ with $KL(a,b) = 0$ iff $a = b$, but it is not formally a distance measure as it does not obey the triangle inequality. It turns out that for high $\delta$ the KL distance measure is directly linked to our approximation.

**Proposition 3** For $\delta$ high enough, $\alpha_\delta$ is falling in the KL distance between the state contingent signal densities given a statically optimal action $a^*(0)$.

**Proof:** Take the derivative of $\alpha_\delta$ in $\delta$ evaluated at $\delta = 1$ we find:

$$\left| \frac{\partial \alpha_\delta}{\partial \delta} \right|_{\delta=1} = 1 + KL(f^H(\cdot, a^*(0)), f^L(\cdot, a^*(0))).$$

Also $\alpha_\delta$ is continuous in $\delta$. Thus, for $\delta$ near 1, higher $KL$ implies $\alpha_\delta$ falls more rapidly as we reduce $\delta$. This in turn implies that a higher $KL$ implies a lower $\alpha_\delta$ for high enough $\delta$. \qed
Notice that our approximation for $v$ involves three terms. The linear portion is only a function of the flow profit function. For $x$ close enough to 0, the final term will be larger iff $\alpha_\delta$ is smaller, regardless of the value of $C$. This in turn implies that for $x$ low enough, changes in the signal structure will increase the value function, iff they increase the KL distance $KL(f^H(\cdot, a^*(0)), f^L(\cdot, a^*(0)))$.

6 Applications

Now that we have a sharp characterization of the value function in the neighborhood of a fixed point we can use this in some applications. First we use our result to provide a lower bound on $\delta$ such that complete learning obtains. We then ask what the implications of our approximation are for the value of a small signal.

6.1 Complete Learning

Proposition 2 establishes the behavior of the value function near any fixed point $x^*$ for all $\delta$ above a given threshold. However, it turns out that our approximation only holds for $x^* \in \{0, 1\}$. The reason? For an interior fixed point to exist two criterion must be met: there must exist an interior belief such that the statically optimal action given that belief results in no learning, and this statically optimal action must be dynamically optimal. Thus, at any interior fixed point there is a tradeoff between maximizing static profits in the current period and acquiring information with which to increase flow payoffs in the future.

As $\delta$ increases, two things happen: the future becomes more important relative to the current period, but also the value function flattens out (becomes less convex). Easley and Kiefer (1988) showed that for $\delta$ high enough the former effect must dominate and there cannot exist an interior fixed point. We provide some insight into this result by using our asymptotic expression to explicitly calculate the static losses and dynamic gains from a marginal change in the statically optimal policy. We then show that our functional form implies that the dynamic gains of such a change must swamp the static losses.
Proposition 4 (Easley-Kiefer (1988)) If $f^\theta$ is $C^2$ then our asymptotic approximation to the value function is inconsistent with interior fixed points $x^* \in (0, 1)$, when $\alpha_\delta \in (1, 2)$, i.e. for $\delta > \delta^*$

In Appendix A we prove this by contradiction: first assume that an interior $x^*$ is a fixed point. Then use our approximation to the value function in the neighborhood of this fixed point. Then show that an $\epsilon$ deviation from the optimal policy increases the total expected payoff; thus showing that the policy could not have been optimal.

The intuition for the result is straightforward. As mentioned above, at an interior fixed point the statically optimal policy is chosen and no information is revealed. By the Envelope Theorem, deviating from the statically optimal policy by some small amount $\epsilon$, results in gains proportional to some bounded constant times $\epsilon^2$. However the informational gains are roughly proportional to $\epsilon^2$ times the second derivative of the value function (assuming for the sake of this intuition only that such a second derivative exists). But we have shown that this second derivative must be unbounded as $x \to x^*$ for $\delta$ above our computed threshold. Thus, the gains of a marginal change from the statically optimal policy swamp the losses for small enough deviations.

**Example: The Ignorant Monopolist Revisited.** As discussed above, the monopolist can only converge to an interior fixed point when conditional demand curves cross at some price $\hat{a}$, and the statically optimal policy is $a^0(x^*) = \hat{a}$ for some $x^* \in (0, 1)$. Thus, one can calculate all possible interior fixed points from the static optimization problem alone. One can then calculate a $\delta^*$ threshold for all such prices $\hat{a}$ where conditional expected demands cross using our formula, as follows:

$$
1/\delta^* \equiv \Phi(2, x^*) = \lim_{x \to x^*} \left[ \frac{(f^H(\hat{a})x - x^*f(x, \hat{a}))^2}{f(x, \hat{a})(x - x^*)^2} + \frac{(1 - f^H(\hat{a}))x - x^*(1 - f(x, \hat{a}))^2}{(1 - f(x, \hat{a}))(x - x^*)^2} \right]
$$

Then $\forall \delta > \delta^*$, $x^*$ cannot be the fixed point of a dynamically optimal policy.

### 6.2 The Value of a Signal

Assume the decision maker facing our learning problem is given the chance to observe an extra signal with densities $g^\theta(\sigma, \varepsilon)$ for $\theta \in \{H, L\}$, where $\varepsilon \in \mathbb{R}_+$. Assume $g^\theta$ is $C^2$ in $\varepsilon$ and that $g^H(\sigma, \varepsilon) \neq g^L(\sigma, \varepsilon)$ on a set of positive measure for all $\varepsilon > 0$, and
\(g^H(\sigma, 0) = g^L(\sigma, 0) \forall \sigma\). Thus, we have an informative signal for all non zero \(\varepsilon\), but the signal becomes uninformative as \(\varepsilon \rightarrow 0\). Let \(g(\sigma, x, \varepsilon) = xy^H(\sigma, \varepsilon) + (1 - x)y^L(\sigma, \varepsilon)\). Overuse notation and let \(z(\sigma, x, \varepsilon)\) be the updated value of \(x\) upon observing signal \(\sigma\) generated by \(g\), so that \(z(\sigma, x, \varepsilon) = (g^H(\sigma, \varepsilon)x)/g(\sigma, x, \varepsilon)\). Having shown that interior fixed points cannot exist when our approximation holds, we focus on \(x^* \in \{0, 1\}\). In fact, we restrict our analysis to \(x^* = 0\) as the analysis in the \(x^* = 1\) case is identical.

Given \(\varepsilon\) and value function \(v\) the value of a one time signal is:

\[
\mathcal{V}(x, \varepsilon) \equiv \int v(z(\sigma, x, \varepsilon))g(\sigma, x, \varepsilon)d\sigma - v(x)
\]

We ask, what is the value of the signal parameterized by \(\varepsilon\) for \(x\) near \(x^*\). Clearly, this value vanishes as \(x \rightarrow x^*\). Thus, we must normalize appropriately. It turns out that \(\mathcal{V}(x, \varepsilon)/x^\alpha\delta\) is the appropriate normalization to study.

**Proposition 5** Let the premise of Proposition 2 hold, and assume \(\delta > \delta^*\). Then:

\[
\lim_{x \rightarrow x^*} \frac{\mathcal{V}(x, \varepsilon)}{x^{\alpha\delta}} = c \left[ \int \frac{g^H(\sigma, \varepsilon)^{\alpha\delta}}{g^L(\sigma, \varepsilon)^{\alpha\delta-1}}d\sigma - 1 \right]
\]

**Proof:** If our approximation is valid, then we have:

\[
\lim_{x \rightarrow 0} \mathcal{V}(x, \varepsilon) = c \lim_{x \rightarrow 0} \left[ \int \left( \frac{z(\sigma, x, \varepsilon)}{x} \right)^{\alpha\delta} g(\sigma, x, \varepsilon)d\sigma - 1 \right] = \lim_{x \rightarrow 0} \left[ \int \left( \frac{g^H(\sigma, \varepsilon)}{g(\sigma, x, \varepsilon)} \right)^{\alpha\delta} g(\sigma, x, \varepsilon)d\sigma - 1 \right] = c \left[ \int \frac{g^H(\sigma, \varepsilon)^{\alpha\delta}}{g^L(\sigma, \varepsilon)^{\alpha\delta-1}}d\sigma - 1 \right]
\]

Note that changes in \(\alpha\delta\) impact the value of a marginal signal in two ways: via the exponent and via the coefficient \(\alpha\delta(\alpha\delta - 1)\), thus changes in the information structure that increase \(\alpha\delta\) may increase or decrease the value of a marginal signal depending on the value of \(x\). However, for low or high enough \(x\) the exponent dominates and increasing \(\alpha\delta\) decreases the value of a marginal signal. Thus, \(\alpha\delta\) fully orders the value of a marginal signal for decision makers who are almost sure about the state of the world.

23
A Proof of Proposition 4

Let $x^* \in (0, 1)$ be a fixed point of a dynamically optimal policy $a^*$, and consider an $\varepsilon$ deviation from $a^*$ at $x = x^*$. If we can show this increases the discounted value as implied by our approximation to $v$ then we are done. To simplify notation, define:

$$\hat{z}(\sigma, \varepsilon) \equiv z(\sigma, x^*, a^*(x^*) + \varepsilon) \quad \hat{f}^\theta(\sigma, \varepsilon) = f^\theta(\sigma, a^*(x^*) + \varepsilon) \quad \hat{\pi}(\varepsilon) \equiv \pi(x^*, a^*(x^*) + \varepsilon)$$

and $\hat{f} \equiv x^* f^H + (1 - x^*) f^L$. We need to show that there exists non zero $\varepsilon$ such that:

$$(1 - \delta)[\hat{\pi}(\varepsilon) - \hat{\pi}(0)] + \delta \int v(\hat{z}(\sigma, \varepsilon)) \hat{f}(\sigma, \varepsilon) d\sigma > (1 - \delta)\hat{\pi}(0) + \delta v(x^*)$$

Since $v(\hat{z}(\sigma, 0)) = v(x^*)$ as $x^*$ a fixed point. We established in Claim 1 of the proof of Proposition 2 that there exists a linear function $L(x)$ and $A > 0$ such that $v(x) \geq L(x) + A(x - x^*)a^*$. Further $L(x^*) = v(x^*)$ thus, the following inequality is sufficient:

$$(1 - \delta)[\hat{\pi}(\varepsilon) - \hat{\pi}(0)] + \delta A \int (z(\sigma, \varepsilon) - x^*) a^* \hat{f}(\sigma, \varepsilon) d\sigma$$

Define the static loss as: $L(\varepsilon) \equiv (1 - \delta)[\hat{\pi}(\varepsilon) - \hat{\pi}(0)]$, and the dynamic gain as: $G(\varepsilon) \equiv \delta A \int (z(\sigma, \varepsilon) - x^*) a^* \hat{f}(\sigma, \varepsilon) d\sigma$. To complete the proof we shall establish the following two claims:

Claim 1: $\lim_{\varepsilon \to 0} \frac{L(\varepsilon)}{\varepsilon^2} > -\infty$

Claim 2: $\lim_{\varepsilon \to 0} \frac{G(\varepsilon)}{\varepsilon^a} > 0$

Proof of Claim 1: By Taylor’s Theorem, we have:

$$L(\varepsilon) = \pi_a(x^*, a^*(x^*)) \varepsilon + O(\varepsilon^2)$$

while, by the Envelope Theorem $\pi_a(x^*, a^*(x^*)) = 0$.

Proof of Claim 2: Using the same steps as we did in the Proof of Lemma 1

$$\lim_{\varepsilon \to 0} (\hat{z}(\sigma, \varepsilon) - x^*) / \varepsilon = \frac{x^*}{f(\sigma, x^*, 0)} \lim_{\varepsilon \to 0} \frac{\hat{f}^H(\sigma, \varepsilon) - \hat{f}^L(\sigma, \varepsilon)}{\varepsilon}$$
Since \( x^* \) is an interior fixed point \( \hat{f}^H(\sigma, 0) - \hat{f}^L(\sigma, 0) = 0 \) a.s. so that we must use L’Hopital to get:

\[
\lim_{\varepsilon \to 0} \frac{\hat{f}^H(\sigma, \varepsilon) - \hat{f}^L(\sigma, \varepsilon)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{\hat{f}_{x_\varepsilon}^H(\sigma, \varepsilon) - \hat{f}_{x_\varepsilon}^L(\sigma, \varepsilon)}{\varepsilon}
\]

which is non zero almost surely, by Local Uniqueness. Which implies that:

\[
\lim_{\varepsilon \to 0} \mathcal{G}(\varepsilon) = \lim_{\varepsilon \to 0} = \int \left( \frac{z(\sigma, \varepsilon) - x^*}{\varepsilon} \right)^2 \hat{f}(\sigma, \varepsilon) d\sigma > 0
\]

Since beliefs are a martingale we have that \( \lim_{\varepsilon \to 0} (\hat{z}(\sigma, \varepsilon) - x^*) \hat{f}(\sigma, \varepsilon) d\sigma = 0 \). We can show that \( \lim_{\varepsilon \to 0} [\hat{z}(\sigma, \varepsilon) - x^*] \hat{f}(\sigma, \varepsilon) d\sigma \) is continuous and strictly increasing in \( \alpha \), with like reasoning to that in the proof of Lemma 3. All together, this implies the desired result.

\[\square\]

References


